

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-286889

(43)Date of publication of application : 01.11.1996

(51)Int.Cl.

G06F 5/00

(21)Application number : 07-089030

(71)Applicant : HITACHI LTD

(22)Date of filing : 14.04.1995

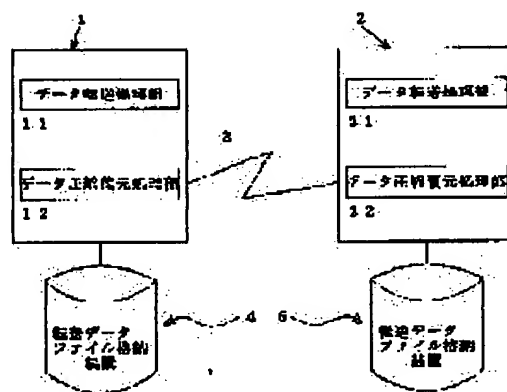
(72)Inventor : FUKUDA MASAJI
SEKI HIROSHI
NAKAYAMA KEIJI

(54) DATA FILE COMPRESSION DECODING METHOD

(57)Abstract:

PURPOSE: To reduce the transfer processing time by compressing data file transfer data based on a data pattern designated at transfer of a data file or a division period.

CONSTITUTION: When the user designates a data pattern or a division period to a data compression decoding processing section 12 in advance, a data file stored in a transfer data file storage device 4 is read, and in the case of designating the data pattern, a pattern is searched to divided data, and in the case of designating the division period, difference data are generated while conducting division for each division period and the generated difference data are given to a data transfer processing section 11. The data transfer processing section 11 sends data to a data transfer processing section 21 through a communication channel. The data transfer processing section 21 gives the received data to a data compression decoding processing section 22. The data compression decoding processing section 22 decodes the received data and generates a file and they are stored in a transfer data file storage device 5.



(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平8-286889

(43)公開日 平成8年(1996)11月1日

(51)Int.Cl.⁶

G 0 6 F 5/00

識別記号

庁内整理番号

F I

G 0 6 F 5/00

技術表示箇所

H

審査請求 未請求 請求項の数1 O L (全 4 頁)

(21)出願番号 特願平7-89030

(22)出願日 平成7年(1995)4月14日

(71)出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72)発明者 福田 正路

神奈川県横浜市戸塚区戸塚町5030番地 株式会社日立製作所ソフトウェア開発本部内

(72)発明者 関 宏

神奈川県横浜市戸塚区戸塚町5030番地 株式会社日立製作所ソフトウェア開発本部内

(72)発明者 中山 敬二

神奈川県横浜市戸塚区戸塚町5030番地 株式会社日立製作所ソフトウェア開発本部内

(74)代理人 弁理士 小川 勝男

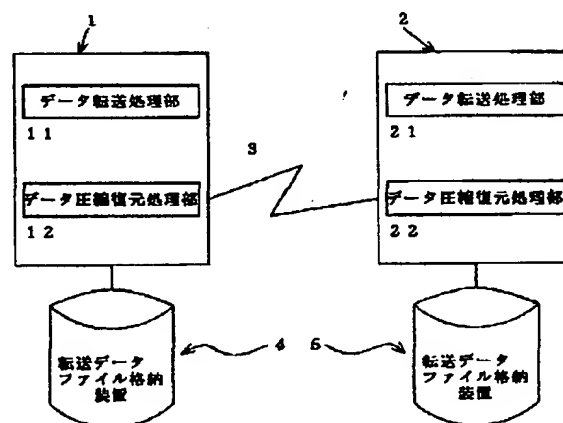
(54)【発明の名称】 データファイル圧縮復元方法

(57)【要約】

【目的】本発明の目的は、データファイル転送時指定したデータパターンか分割周期を元にデータファイル転送データを圧縮し転送処理時間を削減することにある。

【構成】利用者はあらかじめデータパターンか分割周期をデータ圧縮復元処理部12に指定すると転送データファイル格納装置4に格納されたデータファイルを読み込み、データパターン指定の場合はパターンサーチを行ってデータの分割を行い、分割周期指定の場合は分割周期ごとに分割を行いながら差分データの作成を行い、作成した差分データをデータ転送処理部11へ渡す。データ転送処理部11は通信回線を通じてデータ転送処理部21へデータの送信を行う。データ転送処理部21は受信したデータをデータ圧縮復元処理部22へ渡す。データ圧縮復元処理部22は渡されたデータをもとにデータの復元を行いながらファイルの作成を行い転送データファイル格納装置5に格納を行う。

図1



1

【特許請求の範囲】

【請求項1】通信回線により統合されたコンピュータシステム間でデータファイルを送受信するデータファイル圧縮復元方法において、利用者が指定した周期を単位とした差分の抽出により転送データを分割し、該分割した転送データ間の差分データを抽出し、該差分データの長さ、ロケーション及び内容を送信し、受信側で受信したデータを単位間の差分データをもとに復元することを特徴とするデータファイル圧縮復元方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、コンピュータシステム間で送受信するデータ量を削減し、処理の効率を向上するための、データファイル圧縮復元方法に関する。

【0002】

【従来の技術】従来のデータファイル圧縮復元方法は、転送データの圧縮は同一文字が連続した場合に文字列のサブレスを行う方法であり、本発明のように利用者が指定したデータパターンや周期により分割したデータ間の差分抽出による圧縮は行っていないかった。

【0003】

【発明が解決しようとする課題】一般のファイルは内容や形式はそれぞれ異なるが、その多くはページ等の大きな単位で同じ形式の繰り返しデータであり、各単位の差分はページ全体より情報量が少なくデータの圧縮が容易に行える。上述したように、従来は転送データの圧縮を行う場合、対象となる繰り返しデータはプログラムのロジックにより動的に決定していた。そのため圧縮処理ロジックが複雑になり圧縮処理時間が大きくなるという課題があった。

【0004】本発明は、従来の技術に内在する上記課題を解決するためになされたものであり、本発明の目的は差分抽出の周期を利用者があらかじめ指定することにより高速で効率の高い簡単なデータファイル圧縮復元方法を提供することにある。

【0005】

【課題を解決するための手段】上記目的を達成するために、本発明に係るデータファイル圧縮復元方法は通信回線で接続された2台以上のデータファイル送受処理系コンピュータシステムにより構成されたデータファイル転送システムにおいて、転送対象データファイルの転送開始時に利用者が指定した分割周期を決定するためのデータパターンまたは分割周期を元に転送データを分割して、分割単位間の差分を抽出して送信する処理部分と、受信するデータを単位間の差分をもとに受信データを復元する処理部分により構成される。

【0006】

【作用】コンピュータシステム1からコンピュータシステム2に対してデータファイルを転送する場合、コンピュータシステム1において利用者がデータパターンを指

2

定した場合、指定データパターンがデータファイル内において出現してから次に出現するまでの間隔を1単位として分割を行い、また、利用者が分割周期を指定した場合はその周期によりデータファイルの分割を行う。これら分割処理を行いながら、最初に先頭の単位のみを送信を行い、その後は前に送信した単位との差分のみを順次送信する。コンピュータシステム2においてコンピュータシステム1より先頭の分割単位を受信した後に受信する差分データをもとに先頭以降の分割単位の復元を行ないながら転送データファイル全体の復元処理を行なう。

【0007】

【実施例】本発明の一実施例について図面に基づいて具体的に説明する。

【0008】図1は本発明の一実施例を示すデータファイル圧縮復元方法の概念を示す構成図である。1及び2はコンピュータネットワーク上のコンピュータシステム、3はネットワーク内のコンピュータシステム間を接続する通信回線、4及び5は転送または受信したデータファイルを格納する転送データファイル格納装置を示す。また、コンピュータシステム1内において、11はコンピュータシステム1内で作成された差分データをコンピュータシステム2へのデータ転送を行なうデータ転送処理部、12は利用者が指定したデータパターンまたは分割周期を元にした差分データのデータ圧縮を行なうデータ圧縮復元処理部であり、コンピュータシステム2内において、21はコンピュータシステム1から送られたデータの受信を行なうデータ転送処理部、22は差分データを元に受信したデータファイルのデータ復元を行なうデータ圧縮復元処理部である。

【0009】図2に示すように「DATA_PAGE=」というデータパターンが一定の周期で現われる内容のデータファイルにおいて、利用者がデータパターンとして「DATA_PAGE=」を指定した場合、データ圧縮復元処理部12は転送データファイル格納装置4内のデータファイルの内容を読み込み、データパターン「DATA_PAGE=」が出現する単位に分割を行う。

【0010】また、図3に示すように利用者が分割周期をLとした場合、データ圧縮復元処理部12は転送データファイル格納装置4内のデータファイルの内容を読み込み、長さLの単位に分割を行う。

【0011】これら図2及び図3で示した2つの方法にて分割したデータの内容の例を図4に示す。データ圧縮復元処理部12は分割単位1をデータ転送処理部11へ渡し、分割単位1と分割単位2との差分をパターンサーチにより求め、先頭の送信単位と次送信単位との差分を図5に示すように差分データの長さや転送単位の先頭からのロケーション及び差分データの内容からなる形式に編集作成する。

【0012】その後、同様に差分データからなる送信単位を作成しながら順次送信単位にデータをデータ転送処

3

理部11へ渡す。データ転送処理部11は受け取った転送単位をコンピュータシステム2に対して転送を行う。

【0013】コンピュータシステム2においてデータ転送処理部21は受信したデータを順にデータ圧縮復元処理部22に渡す。データ圧縮復元処理部22は図5に示す内容のデータを受け取り、先頭の送信単位をそのまま最初の分割単位とし、その後は先頭の送信単位の長さの次に受信する送信単位の内容を解析しながら順次データの復元処理を行う。

【0014】図5に示す送信単位2から図4に示す分割単位2を復元する場合、はじめに受信した送信単位1に対して送信単位2の情報から送信単位1の相対14の位置にあり長さ1のデータである「1」を「2」に置き換える。同様にして差分データ2と差分データ3についての置き換えを行い、これを分割データ2とする。次の分割単位3は分割単位2を元にして送信単位3の差分データの情報から復元を行う。このようにして復元した各分割単位を連結して作成したデータファイルを転送データファイル格納装置5に格納する。

【0015】

【発明の効果】以上説明したように、本発明は利用者が指定したデータパターンまたは分割周期により分割したデータ間の差分情報のみを転送することによりデータフ

4

ァイルの圧縮復元処理を効率よく高速に行うことができるという効果が得られる。

【図面の簡単な説明】

【図1】本発明の一実施例を示すデータファイル圧縮復元方法の概念を示す構成図である。

【図2】転送データファイルを利用者が指定したデータパターンにより分割したデータ内容を示す図である。

【図3】転送データファイルを利用者が指定した分割周期により分割したデータ内容を示す図である。

【図4】分割後のデータ内の差分データの内容を示す図である。

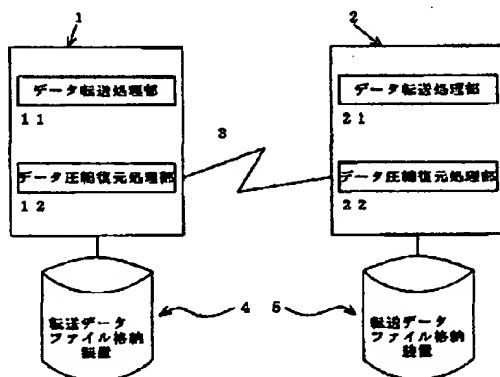
【図5】圧縮処理を行った後の転送データの内容を示す図である。

【符号の説明】

- 1…コンピュータシステム（送信側）
- 2…コンピュータシステム（受信側）
- 11…データ転送処理部（送信側）
- 21…データ転送処理部（受信側）
- 12…データ圧縮復元処理部（送信側）
- 22…データ圧縮復元処理部（受信側）
- 3…通信回線
- 4…転送データファイル格納装置（送信側）
- 5…転送データファイル格納装置（受信側）

【図1】

図1

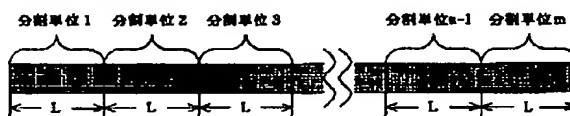


【図3】

図3

分割周期による分割方法

・利用 分割周期をLとした場合



【図2】

図2

データパターンによる分割方法

・利用者がデータパターン＝「DATA_PAGE_」を指定した場合

分割単位1

DATA_PAGE_

.....

分割単位m

DATA_PAGE_

【図4】

図4

分割後のデータ列

分割単位1

DATA_PAGE = 1 DATA1 = 0000_0000 DATA2 = 0100_0000 DATA3 = 0000_0000

分割単位2

差分データ1(長さ=1) 差分データ2(長さ=6) 差分データ3(長さ=3)

DATA_PAGE = 2 DATA1 = 0000_0000 DATA2 = 0011_1100 DATA3 = 0000_FFFF

14 43 68

.....

分割単位79

DATA_PAGE = 79 DATA1 = FFFF_1111 DATA2 = 0000_FFFF DATA3 = 1000_1111

分割単位80

差分データ4(長さ=2) 差分データ5(長さ=1)

DATA_PAGE = 80 DATA1 = FFFF_1111 DATA2 = 0000_FFFF DATA3 = 1000_1111

13 27

↑

内容比較

↑

内容比較

【図5】

図5

圧縮後の送信データの内容例

送信単位1

DATA_PAGE = 1 DATA1 = 0000_0000 DATA2 = 0100_0000 DATA3 = 0000_0000

送信単位2

0104004011110000FFFF

↑ 差分データ3の内容
↑ 差分データ3の先頭からのロケーション
↑ 差分データ3の長さ
↑ 差分データ2の内容
↑ 差分データ2の先頭からのロケーション
↑ 差分データ2の長さ
↑ 差分データ1の内容
↑ 差分データ1の先頭からのロケーション
↑ 差分データ1の長さ

.....

送信単位80

0104004011110000FFFF

↑ 差分データ5の内容
↑ 差分データ5の先頭からのロケーション
↑ 差分データ5の長さ
↑ 差分データ4の内容
↑ 差分データ4の先頭からのロケーション
↑ 差分データ4の長さ